創想館　14-201

2008-4

Tuesday　SPRING

13:00－14:30

2008年度

「情報工学」

第4回講義

前回学んだこと

復 習

- 覚えていなければいけないこと
- 答えられなければならないこと

✓ ‘シャノンの情報量’ および ‘エントロピー’ の式で，対数の底が「2」になっている理由は？

$$H(X) = -\sum_{i=1}^{N} p(a_i)\log_{2} p(a_i) \quad [bit]$$

✓ エントロピーの最大値は？

$$H \le log_2 N \qquad where\ N：事象数$$

〔その証明〕

（1）微分によって，極値を求める． 本日

（2）式の変形によって，恒常的に $H \le log_2 N$ が成立することを示す． 済ませた！

ANIMATION

証明（１）

・・・ 厳密に言うと、証明ではない！

- 解析的に，$H(X)$ の最大値を求める．
- ただし，一般的には無理．
- 事象が３つしかない場合のエトロピーなら，比較的簡単に求まる．

これをやってみよう！

まず， 条件式 ： $\sum p_i = p_1 + p_2 + p_3 = 1$

上式の条件下で，H の最大化を図る．

$$H = -\sum_{i=1}^{3} p_i log_2 p_i = -p_1 log_2 p_1 - p_2 log_2 p_2 - p_3 log_2 p_3$$

まず，p_3を固定（定数化）し，p_1、p_2 を変数とする．

$$p_1 + p_2 = 1 - p_3 = const.$$

ANIMATION

証明（１）　・・・　つづき

まず，$\dfrac{dH}{dp_1}=0$　となるためには．．．

$$\frac{dH}{dp_1}=\frac{dH}{\partial p_1}\frac{\partial p_1}{dp_1}+\frac{dH}{\partial p_2}\frac{\partial p_2}{dp_1}$$　であり

$p_1+p_2+p_3=1$

$$\frac{\partial p_1}{dp_1}=\frac{\partial p_1}{\partial p_1}=1,\quad \frac{\partial p_2}{dp_1}=\frac{\partial p_2}{\partial p_1}=\frac{\partial}{\partial p_1}\left(1-p_3-p_1\right)=-1$$　であるから

$$\frac{\partial H}{\partial p_1}=-\frac{\partial}{\partial p_1}\left(p_1 log_2 p_1\right)=-log_2 p_1-p_1\frac{1}{p_1}=-log_2 p_1-1$$

同様に　$$\frac{\partial H}{\partial p_2}=-\frac{\partial}{\partial p_2}\left(p_2 log_2 p_2\right)=\cdot\ \cdot\ \cdot\ \cdot\ \cdot=-log_2 p_2-1$$

証明（１）　・・・　つづき

$$\frac{dH}{dp_1}=-log_2 p1-1-\left(-log_2 p_2-1\right)$$

$$=-log_2 p_1+log_2 p_2\Rightarrow 0$$

この式が成立するためには．．．

$p_1=p_2$　とならねばならない！　$\because\ p_1+p_2+p_3=1$

同様に，p_2を固定すると　．．．　$p_3=p_1$

P_1を固定すると　．．．　$p_2=p_3$　となるはず．

すなわち，H_{max}　を与えるのは，　$p_1=p_2=p_3=\dfrac{1}{3}$

証明（1） ・・・ つづき

以上から類推するに，N 事象ならば，

$$H_{max} = \log_2 N$$

$$\therefore \quad H \le \log_2 N$$

言葉の定義

冗長度 γ ： $\gamma = \dfrac{H_{max} - H}{H_{max}} = 1 - \dfrac{H}{H_{max}}$

example. X 地方 $H = 1.8$

W 地方 $H_W = 2 = H_{max}$

$\gamma = 1 - 1.8/2 = 0.1$

以上から類推するに，N 事象ならば，

$$H_{max} = \log_2 N$$

$$\therefore \quad H \le \log_2 N$$

‘無駄さ’ の度合

必ず覚えておいて下さい、忘れると試験に失敗します！

言葉の定義

冗長度 γ ： $\gamma = \dfrac{H_{max} - H}{H_{max}} = 1 - \dfrac{H}{H_{max}}$

example. X 地方 $H = 1.8$

W 地方 $H_W = 2 = H_{max}$

$\gamma = 1 - 1.8/2 = 0.1$

以上が

いて
ると
ます！

言葉の

前出の例題（天気予報）の場合は・・・

天気＼地方	X	Y	Z	W
F	1/2	7/8	1	1/4
C	1/4	0	0	1/4
R	1/8	1/8	0	1/4
S	1/8	0	0	1/4
γ	0.1	0.75	1	0

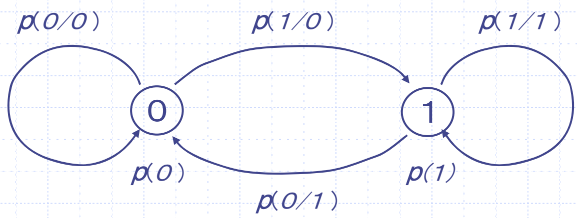
天気予報は不要！

天気予報は有意義！

$\gamma = 1 - 1.8/2 = 0.1$

§ 3.2 マルコフ過程のエントロピー

バイナリー符号系列の単純マルコフ過程



マルコフ過程で考える（遷移確率を加味する）と
エントロピーは，どうなるか？

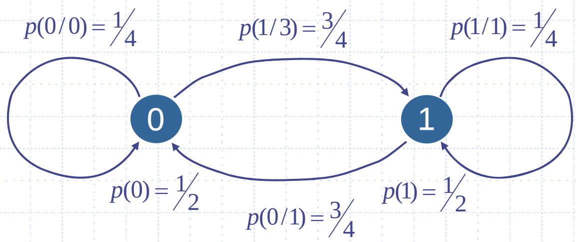
［定義］　単純マルコフ過程のエントロピー

$$H_{M1} = -\sum_{i,j=1}^{N} p(i)\,p(j/i)\log_2 p(j/i) \quad (bit)$$

$$cf.\ \ H = -\sum_{i=1}^{N} p(i)\log_2 p(i)$$

Example 3.2.1

０，１の符号系列（０１１００１０１１００１０１１１…………）を，
単純マルコフ過程と見た時のエントロピーを求めよ．
ただし，０，１の生起確率および遷遷移確率は，それぞれ
$p(0)=p(1)=1/2$、$p(0/0)=p(1/1)=1/4$、$p(1/0)=p(0/1)=3/4$ とする．
また，同じ符号系列を，ただの確率過程と見た時のエントロピーはどうなるか？



略解：

$$H_{M1} = -\underbrace{p(0)}_{1/2}\underbrace{p(0/0)}_{1/4}/\log_2 \underbrace{p(0/0)}_{1/4} - \underbrace{p(0)}_{1/2}\underbrace{p(1/0)}_{3/4}\log_2 \underbrace{p(1/0)}_{3/4}$$

$$-\underbrace{p(1)}_{1/2}\underbrace{p(1/1)}_{1/4}\log_2 \underbrace{p(1/1)}_{1/4} - \underbrace{p(1)}_{1/2}\underbrace{p(0/1)}_{3/4}\log_2 \underbrace{p(0/1)}_{3/4}$$

$$= . \ . \ . \ . \ . \ . \ . \ . = \underline{0.81\,bit}$$

どう見るか？

単純な確率過程（マルコフ過程ではない）と見ると，

$$H_M = -\underbrace{p(0)}_{1/2}\log_2 \underbrace{p(0)}_{1/2} - \underbrace{p(1)}_{1/2}\log_2 \underbrace{p(1)}_{1/2} = \underline{1\ bit}$$

よって、 $H_{M1} \leq H$

このことは，一般的にいえる．

すなわち，マルコフ過程で考えると，エントロピーは常に小さくなる！

略解：

$H_{M1} = -p(0)\,p(0/0)/\log_2 p(0/0) - p(0)\,p(1/0)\log_2 p(1/0)$

1/2 1/4 1/4 3/4 3/4

$- p(1)\,p(1/$ … $p(1/1)$ … $0/1)\log$ … $p(0/1)$

3/4

マルコフ過程で考えると，

エントロピーが小さくなる！

どう見るか？

常に

単純な確…

H_M … $= 1\ bit$

よって、 … $_{M1} \le H$

このことは，一般的にい…る.

すなわち，マルコフ過程で考えると，エントロピーは常に小さくなる！

このことを，気温の変化を例にとって考えてみよう…

- 気温は，日々，突然（ランダムに）変化しているわけではない.

⇒ 昨日は真夏日で，今日は氷点下などということは，まずない！

- 昨日の温度と今日の温度の間には，何がしかの関係がある.

⇒ 昨日と今日の最高気温に10℃の差があれば，ニュースになる！

ANIMATION

このことを，気温の変化を例にとって考えてみよう…

- 気温は，日々，突然（ランダムに）変化しているわけではない．

⇨ 昨日は真夏日 ない！

> 今年は，ゴールデンウィーク中に25℃を超える日が何日かあった．
> ところが，その直後（ここ数日）3月中旬の気温が続いている．だから，それがニュースになっている．

- 昨日の温度と今日の 係がある．

⇨ 昨日と今日の最高気温に10℃の差があれば，ニュースになる！

- よって，「一般に、事象の時間的推移を考慮すると，エントロピーは低下する」と言える．

ANIMATION

これで，「易しい話」は、一段落！

ここから，段々難しくなる！

まずは，通報に「誤り」があったり，
「嘘」があった場合に，
エントロピーは，どうなるか？・・・
・・・と，いう話からいく！

ANIMATION

§ 3.3　誤りによる情報量の低減

今回の授業（以下の部分）は，失敗しました！
内容と時間配分が上手くいかず，盛沢山過ぎて，
消化不良をきたしてしまったのではないか
と思います．すみませんでした．
以下の部分については，次回十分に復習しますので，
お許し頂きたく思います．

$$cf.\quad \text{情報量} = log_2 \frac{1}{\text{その事象が生起する確率}}$$

Shannonによる情報量の概念

§ 3.3　誤りによる情報量の低減

- 発言に，誤りがあったり，嘘があったりすると
その発言者の発言は情報量が低下する．**常識！**

イソップ童話：狼と少年

［定義］誤りがある場合の情報量

$$\left[\begin{array}{c}\text{誤りがある場合}\\\text{の情報量}\end{array}\right] = \left[\begin{array}{c}\text{誤りがない場合}\\\text{の情報量}\end{array}\right] - \left[\begin{array}{c}\text{誤りによって}\\\text{失われる情報量}\end{array}\right]$$

$$\left[\begin{array}{c}\text{誤りによって失われる}\\\text{情報量}\end{array}\right] = log_2 \frac{1}{\text{その事象が本当に生起する確率}}$$

$$cf.\quad \text{情報量} = log_2 \frac{1}{\text{その事象が生起する確率}}$$

Shannonによる情報量の概念

§3.3 誤りによる情報量の低減

▼ 発言… そ…

この気持ち・・・
解るか？

［定義］誤りが…

$$\begin{bmatrix}\text{誤りがある場合}\\ \text{の情報量}\end{bmatrix} = \begin{bmatrix}\text{誤りがない場合}\\ \text{の情報量}\end{bmatrix} - \begin{bmatrix}\text{誤りによって}\\ \text{失われる情報量}\end{bmatrix}$$

$$\begin{bmatrix}\text{誤りによって失われる}\\ \text{情報量}\end{bmatrix} = log_2 \frac{1}{\text{その事象が本当に生起する確率}}$$

$$cf. \quad \text{情報量} = log_2 \frac{1}{\text{その事象が生起する確率}}$$

Shannonによる情報量の概念

誤りがある場合のエントロピー：

・送信側から見たエントロピーという意味
・送信エントロピー＝受信エントロピー
・詳しくは、後で！

$$\begin{bmatrix}\text{誤りがある場合の}\\ \boxed{\text{送信エントロピー}}\end{bmatrix} = \begin{bmatrix}\text{誤りがない場合の}\\ \text{送信エントロピー}\end{bmatrix} - \begin{bmatrix}\text{誤りによって失われる}\\ \text{送信エントロピー}\end{bmatrix}$$

$$\Rightarrow \quad H' = H(x) - H(x/y)$$

この式の意味するところは、

送信通報

$x_1, x_2 \cdot\cdot\cdot x_i, x_N$

$H(x)$

通信路

受信通報

$y_1, y_2 \cdot\cdot\cdot y_i, y_N$

$H(y)$

ANIMATION

誤りがある場合のエントロピー：

$$\begin{bmatrix}\text{誤りがある場合の}\\ \text{送信エントロピー}\end{bmatrix} = \begin{bmatrix}\text{誤りがない場合の}\\ \text{送信エントロピー}\end{bmatrix} - \begin{bmatrix}\text{誤りによって失われる}\\ \text{送信エントロピー}\end{bmatrix}$$

$$\Rightarrow \boxed{H' = H(x) - H(x/y)}$$

情報の消散

この式の意味するところは，

送信通報

$x_1, x_2 \cdot\cdot\cdot x_i, x_N$

$H(x) - H(x/y)$

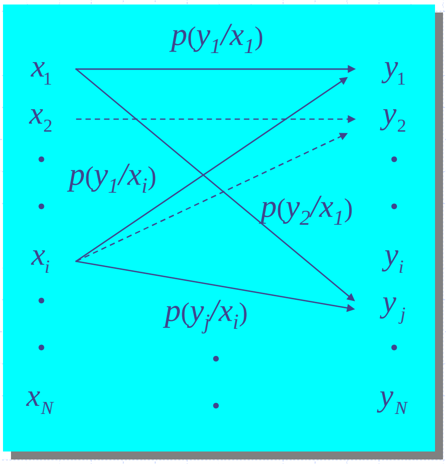
通信路

誤りの混入

受信通報

$y_1, y_2 \cdot\cdot\cdot y_i, y_N$

$H(y) - H(y/x)$



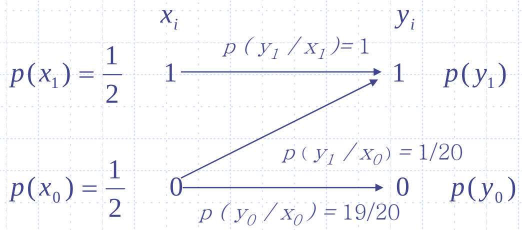
$$H(x) = -\sum_{i=1}^{N} p(x_i)\log_2 p(x_i)$$

必ず，
覚えておくこと！

$$H(x/y) = -\sum_{i=1}^{N} p(y_i)\,p(x_j/y_i)\log_2 p(x_j/y_i)$$ あいまい度

Example 3.3.1

各通報の生起確率と送信誤り率（遷移確率）が，下図によって与えられる１、０の通報系列に対し，送信エントロピーを算出せよ．



送信通報 x_i : 1 0 1 0 1 0 0 1 0 1 1 1 0 1 0 0 ・・・・・

誤　　　　誤

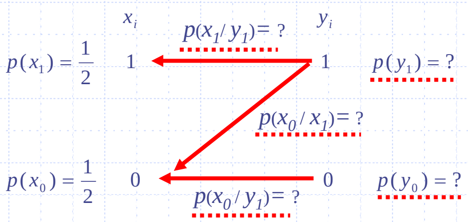
受信通報 y_i : 1 0 1 1 1 0 0 1 0 1 1 1 0 1 0 1 ・・・・・

[略解]

$$H(x)=-\sum_{i=1}^{2}p(x_i)\log_2 p(xi)=-\frac{1}{2}\log_2\frac{1}{2}-\frac{1}{2}\log_2\frac{1}{2}=1\ bit$$

$H(x/y)=？？？$　　　これを求めるのが簡単でない！

前頁の図は，送信側から見たものであった．
$H(\,x\,/\,y\,)$ を求めるには，受信側から見た図が必要



［略解］ ・・・ つづき

$$p(y_1) = \underbrace{p(x_1)}_{1/2}\underbrace{p(y_1/x_1)}_{1} + \underbrace{p(x_0)}_{1/2}\underbrace{p(y_1/x_0)}_{5/100} = \frac{105}{200}$$

同様にして

$$p(y_0) = p(x_1)p(y_0/x_1) + p(x_0)p(y_0/x_0) = \frac{95}{200}$$

また,

$$\underbrace{p(x_1)}_{1/2}\underbrace{p(y_1/x_1)}_{1} = \underbrace{p(y_1)}_{105/200}\underline{p(x_1/y_1)} \longrightarrow p(x_1/y_1) = 0.952$$

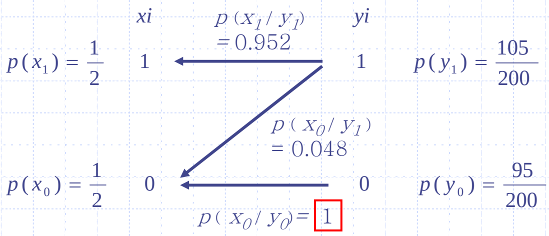
$$\underbrace{p(x_1)}_{1/2}\underbrace{p(y_0/x_1)}_{0} = \underbrace{p(y_0)}_{95/200}\underline{p(x_1/y_0)} \longrightarrow p(x_1/y_0) = 0$$

$$\underbrace{p(x_0)}_{1/2}\underbrace{p(y_1/x_0)}_{5/100} = \underbrace{p(y_1)}_{105/200}\underline{p(x_0/y_1)} \longrightarrow p(x_0/y_1) = 0.048$$

$$\underbrace{p(x_0)}_{1/2}\underbrace{p(y_0/x_0)}_{95/100} = \underbrace{p(y_0)}_{95/200}\underline{p(x_0/y_0)} \longrightarrow p(x_0/y_0) = 1$$

ANIMATION

［略解］ ・・・ つづき



・・・・・ 受信側から観察した図

気付いて欲しいこと ・・・

・ 送信側で観察した場合には，「１」が信頼できる通報であった．

・ しかし，受信側で観察した場合には，「０」が信頼でき通報になっている．

ANIMATION

［略解］ ・・・ つづき

よって，

$$H(x/y) = -\underset{105/200}{p(y_1)}\,\underset{0.952}{p(x_1/y_1)}\log_2 \underset{0.952}{p(x_1/y_1)}$$

$$-\underset{105/200}{p(y_1)}\,\underset{0.048}{p(x_0/y_1)}\log_2 \underset{0.048}{p(x_0/y_1)}$$

$$-\underset{95/200}{p(y_0)}\,\underset{1}{p(x_0/y_0)}\log_2 \underset{1}{p(x_0/y_0)}$$

$$-\underset{95/200}{p(y_0)}\,\underset{0}{p(x_1/y_0)}\log_2 \underset{0}{p(x_1/y_0)} = 0.146\ bit$$

よって，

$$H' = H(x) - H(x/y) = 1 - 0.146 = 0.854\ bit$$

ところで，受信エントロピーという概念もあって良いはず！

$$H'' = H(y) - \underline{H(y/x)}$$

覚えておくこと！

こちらは、「散布度」という.

ここで ．．．

$$H(x) \neq H(y) \qquad H(x/y) \neq H(y/x)$$

で，ある ．．． ところが，

必ず， $\underbrace{H(x) - H(x/y)}_{\text{送信側から見たエントロピー}} = \underbrace{H(y) - H(y/x)}_{\text{受信側から見たエントロピー}}$ となる！

すなわち、
通信路を通して伝送される符号の１符号当たりの平均情報量は、
送信側から見ても、受信側から見ても同じである.

Example 3.3.2：○×問題の採点基準について

問題数 100 問の ○ × 問題があり，全て正答時の ○ × の割合が 50 問－50 問となることが予め知らされている．P 問正解した受験者には，何点を与えるのが適当か？
情報理論的立場から考えよ！

（考え方）

「正答率 p の受験者が，正解１問当たりに送信する平均情報量が，何ビットになるか？」を考えればよい．

宿題
次回までに，考えてくること！

ANIMATION

2008年度

「情報工学」

第4回講義　おわり